Trabalho Prático 2

# Objetivo

Avaliar o entendimento das bibliotecas pandas e streamlit. O projeto deve ser estruturado de forma a explicar o dado, de forma que ao visualizar a sua aplicação streamlit a pessoa seja capaz de entender o que está sendo apresentado a ela.

# Base de Dados

Vocês irão utilizar a base de dados reais do StackOverflow para o ano de 2021.

A base de dados pode ser baixada por meio deste link. <https://info.stackoverflowsolutions.com/rs/719-EMH-566/images/stack-overflow-developer-survey-2021.zip>

Esta base de dados é parecida com a que nós vimos na aula sobre pandas, então podem utilizar o exemplo daquela aula para se guiarem. A única diferença é que lá utilizamos o ano de 2020.

# Processamento

É extremamente importante que seja realizado um processamento da base de dados.

* Limpeza de dados faltantes para informações relevantes. Lembrem-se que nem sempre eu preciso eliminar dados faltantes de todas as colunas e/ou linhas. Algumas vezes podemos preencher estas informações com valores que faça sentido como por exemplo média, mediana, variância, …
* A aplicação streamlit que vocês vão desenvolver deve estar explicando a tomada de decisão que vocês tomaram para responder determinada pergunta. Use textos no streamlit para explicar os detalhes da tomada de decisão que vocês tomaram.
* Vocês podem usar qualquer biblioteca que vocês sentirem confortáveis para realização do projeto. A única obrigatoriedade é que seja feito utilizando streamlit, pandas e alguma biblioteca de visualização de dados, como por exemplo matplolib, plotly ou seaborn.

# Algumas perguntas

Tentem responder as seguintes perguntas, mas não se limitem a elas. Fiquem a vontade para de fato dessacar os dados e acrescentar as perguntas que vocës quiserem:

1. Porcentagem das pessoas que responderam que se consideram profissionais, não profissionais, estudante, hobby, ....
2. Distribuição das pessoas que responderam por localidade. Qual o país que teve maior participação?
3. Qual a distribuição nível de estudo dos participantes?
4. Qual a distribução de tempo de trabalho para cada tipo de profissional respondido na questão 1.
5. Das pessoas que trabalham profissionalmente:
   1. qual a profissão delas?
   2. qual a escolaridade?
   3. qual o tamanho das empresas de pessoas que trabalham profissionalmente.
6. Média salarial das pessoas que responderam.
7. Pegando os 5 países que mais responderam o questionário, qual é o salário destas pessoas?
8. Qual a porcentagem das pessoas que trabalham com python?
9. Sobre python:
   1. Qual o nível de salário de quem trabalha com python globalmente?
   2. Para o Brasil, qual o nível salarial?
   3. Para os 5 países que mais tiveram participação, qual a média salarial?
10. De todos as pessoas, Qual o sistema operacional utilizado por elas?
11. Das pessoas que trabalham com python, qual a distribuição de sistema operacional utilizado por elas.
12. Qual a média de idade das pessoas que responderam?
13. E em python? Qual a média de idade?

Este é apenas algumas perguntas, fiquem a vontade para aprofundar na base de dados respondendo outras perguntas que vocês acharem interessante.